Biotechnol Lett (2009) 31:759–764
DOI 10.1007/s10529-009-9931-6

ORIGINAL RESEARCH PAPER

Insertion of a heterologous gene construct into a non-functional ORF of the *Streptococcus thermophilus* chromosome

John A. Renye Jr. · George A. Somkuti

Received: 5 December 2008 / Accepted: 6 January 2009 / Published online: 27 January 2009

Abstract An integrative vector was constructed for inserting heterologous genes within a non-functional open reading frame (ORF) on the chromosome of *Streptococcus thermophilus.* The vector, pINTRS, contained a temperature sensitive origin of replication and an erythromycin resistance gene for initial selection in *S. thermophilus*. The region of the vector containing unique cloning sites, for insertion of recombinant genes, was flanked by homologous DNA sequences corresponding to a pseudogene in *S. thermophilus* to facilitate chromosomal integration. The gene encoding green fluorescent protein, regulated by a plasmid borne *hsp* promoter of *S. thermophilus*, was cloned into pINTRS to demonstrate proper functioning of the vector.

Keywords *Streptococcus thermophilus* · Integrative plasmid · Green fluorescent protein

Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

J. A. Renye Jr. (✉) · G. A. Somkuti
Department of Agriculture, Eastern Regional Research Center, Agricultural Research Service,
600 E. Mermaid Lane, Wyndmoor, PA 19038, USA
e-mail: john.renye@ars.usda.gov

Introduction

Streptococcus thermophilus is a lactic acid bacterium (LAB) that plays an essential role in dairy fermentations for the production of yogurt and cheese. Since centuries of consumption of this viable bacterium by humans has not resulted in harmful consequences, it is considered a "generally regarded as safe" (GRAS) organism. Its "food-grade" status, along with recent reports showing it may survive passage through the gastrointestinal tract and qualify as a probiotic (Menard et al. 2005), make it an attractive candidate to serve as a host for the expression of heterologous genes to improve food quality or elicit a beneficial response in humans or animals. Recombinant strains of *S. thermophilus* have been constructed for the expression metabolic regulators, such serine hydroxymethyltransferase (Chaves et al. 2002); antimicrobials including pediocin (Coderre and Somkuti 1999); and milk-derived bioactive peptides with antihypertensive and antimicrobial properties (Renye and Somkuti 2008).

However, there is concern about the safety of recombinant bacteria for use in food production. Of major concern is the presence of additional genes on plasmids, such as antibiotic resistance genes, which are required for cloning procedures but not necessary for the desired function of the recombinant bacterium. There is concern over whether or not the use of these plasmids could contribute to the increasing number of antibiotic resistant bacterial pathogens,

thus there is a need to develop new cloning strategies which would not compromise the GRAS status of the bacterium. Potential 'food-grade' selection markers have been explored including genes for sugar metabolism (Lin et al. 1996), peptidase expression (Leenhouts et al. 1991), nisin resistance (McIntyre and Harlander 1993), and green fluorescent protein (GFP) (Solaiman and Somkuti 1997). Another possibility is the use of an integrative vector where only the gene of interest is inserted within the bacterial chromosome. Mollet et al. (1993) reported the insertion of a heterologous gene within the *lac* operon of *S. thermophilus.* Expression of the gene was regulated by the *lac* promoter and as a result of the integration there was a 10-fold decrease in β-galactosidase activity. In addition, the transformation efficiency of this system was 50 to 300-fold less than that normally reported for *S. thermophilus* due to the use of vectors which could not replicate in gram-positive hosts. Other vector systems shown to work in streptococci relies on the temperature sensitive replicon from pVE6002, a broad-host-range replicon that functions in both Gram-negative and Gram-positive bacteria when grown at a permissible temperature, below 35°C (Maguin et al. 1992). Once the temperature is raised, the replicon no longer functions allowing for a recombination event to occur between cloned DNA present on the vector and homologous sequences within the bacterial genome. This replicon functions properly in *S. thermophilus* using either a single vector (Chaves et al. 2002) or a two vector system (Labarre et al. 2001). In both cases, a single homologous DNA fragment was used to facilitate a Campbell-like integration event resulting in potentially unstable transformants with the entire plasmid inserted within the *S. thermophilus* chromosome.

This report describes a strategy for the construction of an integrative vector for inserting heterologous genes within the *S. thermophilus* chromosome at a locus not essential for normal metabolic functions (Bolotin et al. 2004; Hols et al. 2005) that would also result in the elimination of the antibiotic marker gene. The backbone of the integrative vector was pKS1, a plasmid that contains the pVE6002 thermosensitive replicon and resistance genes for both erythromycin and kanamycin, and which effectively inactivates genes in *Bacillus anthracis* (Shatalin and Neyfakh 2005). The constructed vector, pRSINT, was used to test if the *S. thermophilus* plasmid borne promoter, P_{hsp} (Somkuti and Steinberg 1999), could be used to regulate the expression of a single copy, heterologous gene inserted within the chromosome.

Materials and methods

Bacterial strains and plasmids

The plasmid free host *S. thermophilus* ST128 (laboratory collection) was maintained in TYL medium and grown at 37°C (Somkuti and Steinberg 1986). For visualizing GFP expression in *S. thermophilus,* recombinant strains were grown at 32°C. *Escherichia coli* DH5α was used as a host for subcloning of plasmid constructs, and was propagated in BHI medium (BD Diagnostics, Franklin Lakes, NJ) at 37°C. The *Bacillus anthracis* integrative vector pKS1 was a gift from K. Shatalin (New York University School of Medicine). Construction of the *S. thermophilus* expression vectors pRS2 (Renye and Somkuti 2008) and pG341 Pa (Somkuti and Steinberg 1999) were described previously. When appropriate, media were supplemented with erythromycin (Em) at 150 μg ml^{-1} for selection of *E. coli* transformants, and at 15 μg ml^{-1} for selection of *S. thermophilus* transformants.

DNA cloning procedures

Restriction enzymes and T4 DNA ligase were obtained from New England BioLabs (Beverly, MA). PCR primers (Integrated DNA Technologies, Coralville, IA) used for amplification of the pseudogene targeted for integration and the encoded restriction endonuclease sites (underlined) were: gluN1 NotI 5′-CCCGCGGCCGCGAAATGAAA-TGAATAAGC-3′; gluN2 PstI 5′-CCCCTGCAGGAGTTGGCAATCCACAT-3′; gluC1 KpnI 5′-CCCGGTACCCCAGCTCGTTTCAAAGTT-3′; gluC2 BssHII 5′-CCCGCGCGCGGCAC-CTTCAGCTTTCC-3′. PCR primers for amplification of the Em gene and the recombinant *hsp-gfp* construct were: Erm1 5′-CCTCCCGTTAAATAATAG-3′; Erm2 5′-CGATACCGTTTACGAAAT-3′; Hsp 5′-CCACATACTATGCTTGCC-3′; GFP 5′-TTTGTATAGTTCATCCATGC-3′. PCR amplification protocol was as follows: 5 min at 95°C; followed by 35 cycles of 95°C for 30 s, 55°C for 30 s and 74°C for 1 min per 1 kb DNA; with a final extension at 72°C for 5 min. PCR

products used in cloning procedures were purified with the QIAquick PCR purification kit (Qiagen Inc, Valencia, CA). Plasmid isolation from *E. coli* was carried out using a Qiagen Midi kit, or by alkaline lysis followed by CsCl/ethidium bromide ultracentrifugation (Stougaard and Molin 1981) for *S. thermophilus* transformations. Plasmids were recovered from *S. thermophilus* as described previously (Somkuti and Steinberg 1986). Construction of recombinant plasmids was confirmed by PCR with products run on 1% agarose gels in TAE buffer (0.04 M Tris, 0.02 M acetic acid, 0.001 M EDTA, pH 8.0).

Bacterial transformations and heterologous gene expression

Transformation of freshly prepared *E. coli* competent cells was carried out by a heat-shock method (Sambrook et al. 1989), while *S. thermophilus* ST128 was electrotransformed by a protocol previously reported (Somkuti and Steinberg 1988). GFP expression in *E. coli* was observed in colonies under UV illumination (365 nm) and in *S. thermophilus* cells by fluorescent microscopy using a Nikon inverted microscope (Nikon Instruments, Melville, NY), equipped with a UV fluorescent cube (excitation: 365 nm; emission: >400 nm).

Results and discussion

Construction of ST integrative vector pINTRS and pINTRSHG

The plasmid map for the *S. thermophilus* integrative vector, pINTRS, is shown in Fig. 1a. The vector pKS1, developed for the inactivation of genes in *B. anthracis* (Shatalin and Neyfakh 2005), was used as the backbone for designing the architecture of pINTRS. The integrative nature of pKS1 is dependent on two parameters: (1) the presence of a temperature-sensitive origin of replication (Maguin et al. 1992); and (2) the presence of two homologous DNA sequences cloned on either side of the kanamycin resistance gene. When the host bacterium is grown at a non-permissive temperature the replicon no longer functions, forcing a homologous recombination event to occur between the chromosome and cloned DNA sequences which

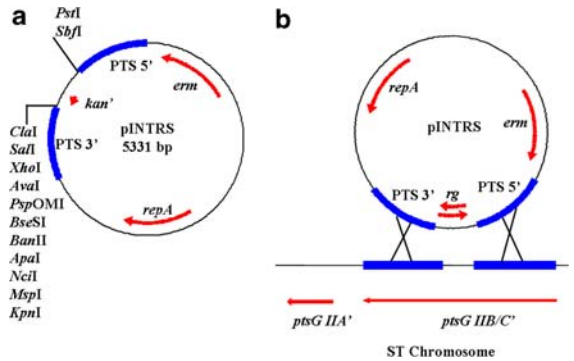

Fig. 1 **a** Plasmid map of pINTRS including the unique restrictions sites for insertion of recombinant promoters and genes. The TS replication gene, *repA*, and erythromycin resistance gene, *erm*, are indicated. The PTS5′ and PTS3′ regions represent cloned DNA fragments corresponding to sequences present in the truncated glucose *ptsIIB/C*. The *kan′* represents the remaining 101 bp from the inactivated kanamycin resistance gene. **b** Integration of a recombinant gene from pINTRS into the *S. thermophilus* chromosome at the locus containing the truncate glucose PTSII ABC genes. On pINTRS, *rg* represents the recombinant gene and the arrows show possible orientations of the gene relative to the truncated glucose *ptsIIB/C*

results in the inactivation of the gene of interest. In the case of pINTRS, the intended purpose of the vector was to deliver recombinant genes into the *S. thermophilus* chromosome for expression from a defined locus that would not interfere with normal metabolic functions of the host and allow its use as a starter culture in dairy fermentations. The chromosomal locus targeted for integration contained pseudogenes encoding truncated components of the glucose phosphoenolpyruvate-dependent phosphotransferase system (PTS) (Fig. 1b) (Bolotin et al. 2004; Hols et al. 2005). The identification of pseudogenes was in agreement with and confirmed previous findings that glucose was a non-PTS sugar and a poor substrate for supporting the growth of *S. thermophilus* (Poolman 1993).

The vector pKS1 (5094 bp) was initially digested with *Eco*RV, removing a 1025 bp fragment containing the first 794 bp of the kanamycin resistance gene, and self-ligated to yield the plasmid pKS1Tr (4069 bp). This step removed the kanamycin gene that was not required for gene inactivation and reduced the overall size of the vector to enable the cloning of large genes or entire operons for expression into the *S. thermophilus* chromosome. DNA fragments required for integration within the glucose PTS locus were amplified by PCR, giving bands of

672 and 654 bp. The 672 bp product contained the 5′-end of the fused *ptsIIB/C* pseudogene and the 654 bp product contained the 3′-end of *ptsIIB/C* (Fig. 1b). The 654 bp product contained a *Kpn*I recognition sequence at its 5′-end and a *Bss*HII sequence at its 3′-end. Following digestion with these two restriction enzymes the DNA fragment was ligated into pKS1Tr at corresponding sites, resulting in the vector pKS1Tr3′. The 672 bp DNA fragment contained a *Not*I recognition sequence at its 5′-end and a *Pst*I recognition sequence at its 3′-end, and was introduced into pKS1Tr3′ at the corresponding sites. The resulting vector, pINTRS, contains the two DNA fragments oriented such that the *ptsIIB/C* pseudogene is encoded on the complimentary strand (Fig. 1a). The DNA sequence between the two cloned fragments contains numerous restriction endonuclease recognition sequences allowing for directional cloning of recombinant genes into the vector (Fig. 1b). The vector was designed without a promoter sequence preset at the insertion locus, thus a promoter must be cloned with the recombinant gene of interest to regulate expression.

To show that the pINTRS functions properly, *gfp* was cloned into the vector under regulation of the *S. thermophilus hsp* promoter (P_{hsp}). GFP expression regulated by P_{hsp} (Somkuti and Steinberg 1999) or P_{2201} (Solaiman and Somkuti 1997) on a shuttle vector has been shown to be readily detectable in *S. thermophilus* colonies when examined under UV light. However, GFP expression regulated by either promoter has not been studied from a single integrated copy of the gene. To test its functionality, the vector pINTRSHG was constructed. The *hsp-gfp* construct was excised from pRS2 (Renye and Somkuti 2008) by digestion with *Hind*III and *Pst*I, releasing a 0.9 kb fragment. The *hsp-gfp* fragment was cloned into pKS1Tr3′ digested with the same restriction enzymes. The *hsp-gfp* fragment had to be cloned into pKS1Tr3′ due to the presence of a *Hind*III restriction sequence presence in the *ptsIIB/C* 5′ DNA fragment. The *ptsIIB/C* 5′ DNA fragment was subsequently cloned into pKS1Tr3′-*hsp-gfp* at the unique *Not*I and *Pst*I restriction sequences, resulting in the vector pINTRSHG with the *hsp-gfp* construct present in the same orientation as *ptsIIB/C* integration fragments (Fig. 2a). Proper construction of the vector was confirmed by the presence of fluorescent *E. coli* DH5α clones (data not shown) and also by the PCR

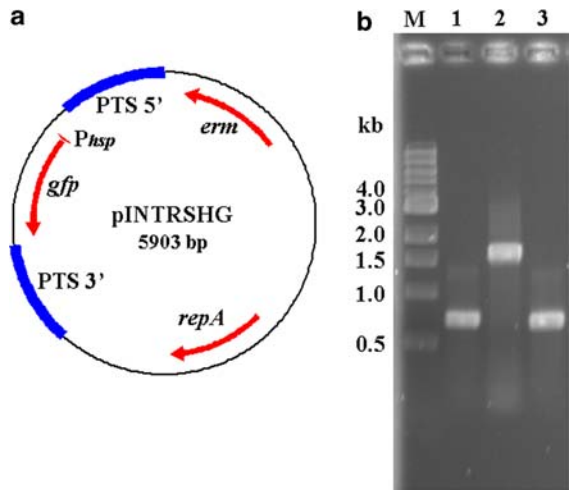

Fig. 2 **a** Map of pINTRSHG. P_{hsp}-*gfp* represents the cloned *S. thermophilus hsp* promoter and the gene encoding GFP. **b** PCR amplification of DNA from pINTRSHG to confirm proper construction of the vector isolated from *E. coli* DH5α. Lane M: 1 kb marker; Lane 1: PTS 5′ region; Lane 2: P_{hsp}-*gfp*-PTS3′; Lane 3: PTS 3′ region

method (Fig. 2b). The *ptsIIB/C* 5′ (Fig. 2b, lane 1) and *ptsIIB/C* 3′ (Fig. 2b, lane 3) fragments were observed in addition to a 1443 bp fragment comprised of *hsp-gfp-ptsIIB/C* 3′ (Fig. 2b, lane 2). Although it has been reported that the pVE6002 replicon is thermosensitive in *E. coli* (Maguin et al. 1992), the pINTRS and pINTRSHG vectors were routinely propagated in *E. coli* DH5α at 37°C under conditions of erythromycin selection.

Insertion of pINTRSHG into *S. thermophilus* ST128

The pINTRSHG was introduced into *S. thermophilus* ST128 by electroporation, with transformants selected for on TYL medium containing erythromycin at 30°C, a permissive temperature for the pVE6002 replicon in Gram-positive hosts (Maguin et al. 1992). Transformants were screened for the presence of pINTRSHG by PCR amplification, showing a 863 bp (lane 1) band and 1570 bp (lane 2) band corresponding to the *hsp-gfp* and *hsp-gfp-ptsIIB/C3′* constructs, respectively (Fig. 3). However, fluorescent colonies were not observed under UV illumination. To ensure that the *hsp-gfp* construct remained functional after introduction into *S. thermophilus*, the pINTRSHG plasmid was recovered from two transformants and reintroduced into *E. coli*

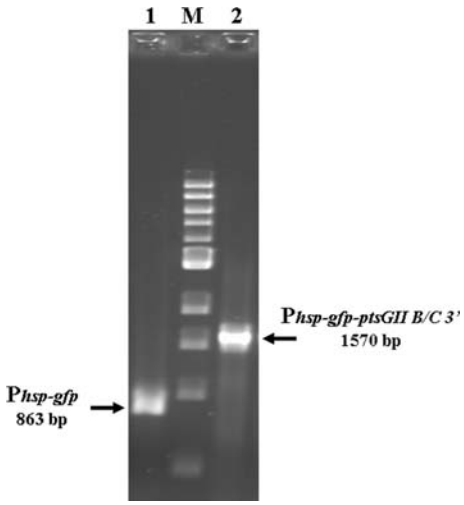

Fig. 3 PCR amplification of regions from pINTRSHG to confirm its presence in *S. thermophilus* transformants maintained at 30°C. Lane 1: Region containing the *hsp* promoter and *gfp*; Lane M: 1 kb marker; Lane 2: Region containing *hsp* promoter, *gfp* and *ptsIIB/C* 3′

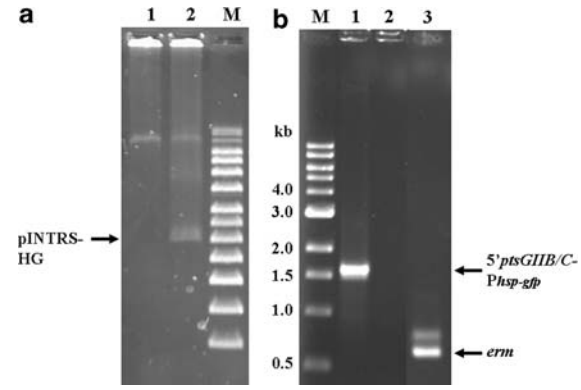

Fig. 4 **a** Plasmid preparation from *S. thermophilus* transformants electroporated with pINTRSHG. Lane 1: Transformants grown at 37°C, a non-permissive temperature for plasmid replication; Lane 2: Transformants maintained at the permissive temperature of 30°C; Lane M: Supercoiled DNA marker. **b** PCR amplification of P_{hsp}-*gfp* construct. Lane M: 1 kb marker; Lane 1: *S. thermophilus* transformants grown at 37°C. **c** PCR amplification of *erm*. *S. thermophilus* transformants grown at 30°C (Lane 1) or at 37°C (Lane 2); Lane M: 1 kb marker

DH5α. The resulting *E. coli* transformants all displayed intense fluorescence under UV illumination. These results suggested that the lack of fluorescence associated with the *S. thermophilus* colonies may be due to a low copy number of pINTRSHG. Previous work showing intense fluorescing *S. thermophilus* colonies containing the same *hsp-gfp* construct were performed using shuttle vectors containing an origin of replication isolated from an endogenous *S. thermophilus* plasmid (Somkuti and Steinberg 1999; Renye and Somkuti 2008). However, single-cell fluorescence was observed when overnight cultures were washed, resuspended in a neutral pH phosphate buffer and examined under a fluorescent microscope, showing that the *hsp-gfp* construct was functional in *S. thermophilus* (data not shown).

Integration of pINTRSHG within the *S. thermophilus* chromosome was accomplished by incubating overnight cultures at the non-permissive temperatures of 37 or 42°C, in the absence of erythromycin selection. The cultures were diluted into fresh TYL and grown overnight at 30°C to facilitate the second homologous recombination event. This was followed by streaking the culture on TYL agar with or without erythromycin. Colonies were observed only in the absence of erythromycin. A culture maintained at the permissive temperature (30°C) was streaked on the same plates as the control, and growth was observed in the presence of erythromycin (data not shown). Fluorescence was observed in individual cells containing the integrated construct under UV illumination. Small scale plasmid preparations were performed on both cultures but plasmid DNA was only recovered from the culture maintained at the permissive temperature of 30°C (Fig. 4a). The presence of the *hsp-gfp* construct in the culture grown at 37°C was confirmed by PCR, which resulted in the amplification of a 1555 bp DNA fragment consisting of the *ptsGIIB/C*-P_{hsp}-*gfp* (Fig. 4b, lane 1). Amplification of *erm* (642 bp) resulted in an intense band for cells maintained at 30°C (Fig. 4b, lane 3) and no band for cells shifted to the non-permissive temperature of 37°C (Fig. 4b, lane 2). This suggested that the initial single crossover event was resolved, thus *erm* was not present in the chromosome. In conclusion, the results show that pINTRS can be used to insert heterologous genes within a non-functional pseudogene on the *S. thermophilus* chromosome and that the plasmid borne *hsp* promoter could regulate expression of a single-copy gene within the chromosome.

Acknowledgments We thank K. Y. Shatalin for providing the integrative vector pKS1, and D. H. Steinberg for technical assistance.

References

Bolotin A, Quinquis B, Renault P et al (2004) Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. Nat Biotechnol 22:1554–1558

Chaves AC, Fernandez M, Lerayer AL et al (2002) Metabolic engineering of acetaldehyde production by *Streptococcus thermophilus*. Appl Environ Microbiol 68:5656–5662

Coderre PE, Somkuti GA (1999) Cloning and expression of the pediocin operon in *Streptococcus thermophilus* and other lactic fermentation bacteria. Curr Microbiol 39:295–301

Hols P, Hancy F, Fontaine L et al (2005) New insights in the molecular biology and physiology of *Streptococcus thermophilus* revealed by comparative genomics. FEMS Microbiol Rev 29:435–463

Labarre C, Schirawski J, van der Zwet A et al (2001) Insertional mutagenesis of an industrial strain of *Streptococcus thermophilus*. FEMS Microbiol Lett 200:85–90

Leenhouts KJ, Kok J, Venema G (1991) Replacement recombination in *Lactococcus lactis*. J Bacteriol 173:4794–4798

Lin MY, Harlander S, Savaiano D (1996) Construction of an integrative food-grade cloning vector for *Lactobacillus acidophilus*. Appl Microbiol Biotechnol 45:484–489

Maguin E, Duwat P, Hege T et al (1992) New thermosensitive plasmid for gram-positive bacteria. J Bacteriol 174:5633–5638

McIntyre DA, Harlander SK (1993) Construction of first-generation lactococcal integrative cloning vectors. Appl Microbiol Biotechnol 40:348–355

Menard S, Laharie D, Asensio C et al (2005) *Bifidobacterium breve* and *Streptococcus thermophilus* secretion products enhance T helper 1 immune response and intestinal barrier in mice. Exp Biol Med (Maywood) 230:749–756

Mollet B, Knol J, Poolman B et al (1993) Directed genomic integration, gene replacement, and integrative gene expression in *Streptococcus thermophilus*. J Bacteriol 175:4315–4324

Poolman B (1993) Energy transduction in lactic acid bacteria. FEMS Microbiol Rev 12:125–147

Renye JA, Somkuti GA (2008) Cloning of milk-derived bioactive peptides in *Streptococcus thermophilus*. Biotechnol Lett 30:723–730

Sambrook J, Maniatis T, Fritsch EF (1989) Molecular cloning a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, New York

Shatalin KY, Neyfakh AA (2005) Efficient gene inactivation in *Bacillus anthracis*. FEMS Microbiol Lett 245:315–319

Solaiman DK, Somkuti GA (1997) Isolation and characterization of transcription signal sequences from *Streptococcus thermophilus*. Curr Microbiol 34:216–219

Somkuti GA, Steinberg DH (1986) Distribution and analysis of plasmids in *Streptococcus thermophilus*. J Ind Microbiol 1:157–163

Somkuti GA, Steinberg DH (1988) Genetic transformation of *Streptococcus thermophilus* by electroporation. Biochimie 70:579–585

Somkuti GA, Steinberg DH (1999) Promoter activity of the pER341-borne ST(Phsp) in heterologous gene expression in *Escherichia coli* and *Streptococcus thermophilus*. FEMS Microbiol Lett 179:431–436

Stougaard P, Molin S (1981) Vertical dye-buoyant density gradients for rapid analysis and preparation of plasmid DNA. Anal Biochem 118:191–193